

RAW SEQUENCE LISTING

**The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.**

Application Serial Number: 09/721,047A
Source: TFW16
Date Processed by STIC: 3-25-05

ENTERED



IFW16

RAW SEQUENCE LISTING

DATE: 03/25/2005

PATENT APPLICATION: US/09/721,047A

TIME: 07:49:59

Input Set : A:\Ocirs396.app

Output Set: N:\CRF4\03252005\I721047A.raw

```

3 <110> APPLICANT: TUMER, NILGUN E.
4   WANG, PINGER
6 <120> TITLE OF INVENTION: TRANSGENIC PLANTS PRODUCING A PAP II PROTEIN
8 <130> FILE REFERENCE: OCIRS 3.9-060 CONT
10 <140> CURRENT APPLICATION NUMBER: 09/721,047A
11 <141> CURRENT FILING DATE: 2000-11-22
13 <150> PRIOR APPLICATION NUMBER: PCT/US99/11301
14 <151> PRIOR FILING DATE: 1999-05-21
16 <150> PRIOR APPLICATION NUMBER: 60/086,374
17 <151> PRIOR FILING DATE: 1998-05-22
19 <160> NUMBER OF SEQ ID NOS: 21
21 <170> SOFTWARE: PatentIn Ver. 3.3
23 <210> SEQ ID NO: 1
24 <211> LENGTH: 1379
25 <212> TYPE: DNA
26 <213> ORGANISM: Phytolacca americana
28 <220> FEATURE:
29 <221> NAME/KEY: CDS
30 <222> LOCATION: (225)..(1163)
32 <400> SEQUENCE: 1
33 ctatgaagtc ggggtcaaagc atatacaggc tatgcattgt tagaaacatt gatgcctctg 60
35 atccccgataa acaatacaaaa ttagacaata agatgacata caagtaccta aactgtgtat 120
37 ggggggagtga aacctcagct gctaaaaaaa cgttgtaaga aaaaaagaaa gttgtgagtt 180
39 aactacaggg cgaaagtatt ggaactagct agtaggaagg gaag atg aag tcg atg 236
40                                     Met Lys Ser Met
41                                     1
43 ctt gtg gtg aca ata tca ata tgg ctc att ctt gca cca act tca act 284
44 Leu Val Val Thr Ile Ser Ile Trp Leu Ile Leu Ala Pro Thr Ser Thr
45   5          10          15          20
47 tgg gct gtg aat aca atc atc tac aat gtt gga agt acc acc att agc 332
48 Trp Ala Val Asn Thr Ile Ile Tyr Asn Val Gly Ser Thr Thr Ile Ser
49          25          30          35
51 aaa tac gcc act ttt ctg aat gat ctt cgt aat gaa gcg aaa gat cca 380
52 Lys Tyr Ala Thr Phe Leu Asn Asp Leu Arg Asn Glu Ala Lys Asp Pro
53          40          45          50
55 agt tta aaa tgc tat gga ata cca atg ctg ccc aat aca aat aca aat 428
56 Ser Leu Lys Cys Tyr Gly Ile Pro Met Leu Pro Asn Thr Asn Thr Asn
57          55          60          65
59 cca aag tac gtg ttg gtt gag ctc caa ggt tca aat aaa aaa acc atc 476
60 Pro Lys Tyr Val Leu Val Glu Leu Gln Gly Ser Asn Lys Lys Thr Ile
61          70          75          80
63 aca cta atg ctg aga cga aac aat ttg tat gtg atg ggt tat tct gat 524
64 Thr Leu Met Leu Arg Arg Asn Asn Leu Tyr Val Met Gly Tyr Ser Asp

```

RAW SEQUENCE LISTING

DATE: 03/25/2005

PATENT APPLICATION: US/09/721,047A

TIME: 07:49:59

Input Set : A:\Ocirs396.app

Output Set: N:\CRF4\03252005\I721047A.raw

```

65 85          90          95          100
67 ccc ttt gaa acc aat aaa tgt cgt tac cat atc ttt aat gat atc tca 572
68 Pro Phe Glu Thr Asn Lys Cys Arg Tyr His Ile Phe Asn Asp Ile Ser
69          105          110          115
71 ggt act gaa cgc caa gat gta gag act act ctt tgc cca aat gcc aat 620
72 Gly Thr Glu Arg Gln Asp Val Glu Thr Thr Leu Cys Pro Asn Ala Asn
73          120          125          130
75 tct cgt gtt agt aaa aac ata aac ttt gat agt cga tat cca aca ttg 668
76 Ser Arg Val Ser Lys Asn Ile Asn Phe Asp Ser Arg Tyr Pro Thr Leu
77          135          140          145
79 gaa tca aaa gcg gga gta aaa tca aga agt cag gtc caa ctg gga att 716
80 Glu Ser Lys Ala Gly Val Lys Ser Arg Ser Gln Val Gln Leu Gly Ile
81          150          155          160
83 caa ata ctc gac agt aat att gga aag att tct gga gtg atg tca ttc 764
84 Gln Ile Leu Asp Ser Asn Ile Gly Lys Ile Ser Gly Val Met Ser Phe
85 165          170          175          180
87 act gag aaa acc gaa gcc gaa ttc cta ttg gta gcc ata caa atg gta 812
88 Thr Glu Lys Thr Glu Ala Glu Phe Leu Leu Val Ala Ile Gln Met Val
89          185          190          195
91 tca gag gca gca aga ttc aag tac ata gag aat cag gtg aaa act aat 860
92 Ser Glu Ala Ala Arg Phe Lys Tyr Ile Glu Asn Gln Val Lys Thr Asn
93          200          205          210
95 ttt aac aga gca ttc aac cct aat ccc aaa gta ctt aat ttg caa gag 908
96 Phe Asn Arg Ala Phe Asn Pro Asn Pro Lys Val Leu Asn Leu Gln Glu
97          215          220          225
99 aca tgg ggt aag att tca aca gca att cat gat gcc aag aat gga gtt 956
100 Thr Trp Gly Lys Ile Ser Thr Ala Ile His Asp Ala Lys Asn Gly Val
101          230          235          240
103 tta ccc aaa cct ctc gag cta gtg gat gcc agt ggt gcc aag tgg ata 1004
104 Leu Pro Lys Pro Leu Glu Leu Val Asp Ala Ser Gly Ala Lys Trp Ile
105 245          250          255          260
107 gtg ttg aga gtg gat gaa atc aag cct gat gta gca ctc tta aac tac 1052
108 Val Leu Arg Val Asp Glu Ile Lys Pro Asp Val Ala Leu Leu Asn Tyr
109          265          270          275
111 gtt ggt ggg agc tgt cag aca act tat aac caa aat gcc atg ttt cct 1100
112 Val Gly Gly Ser Cys Gln Thr Thr Tyr Asn Gln Asn Ala Met Phe Pro
113          280          285          290
115 caa ctt ata atg tct act tat tat aat tac atg gtt aat ctt ggt gat 1148
116 Gln Leu Ile Met Ser Thr Tyr Tyr Asn Tyr Met Val Asn Leu Gly Asp
117          295          300          305
119 cta ttt gaa gga ttc tgatcataaa cataataagg agtatatata tattactcca 1203
120 Leu Phe Glu Gly Phe
121          310
123 actatattat aaagcttaaa taagaggccg tgtaattag tacttggtgc cttttgcttt 1263
125 atggtgttgt ttattatgcc ttgtatgctt gtaattattat ctagagaaca agatgtactg 1323
127 tgtaatagtc ttgtttgaaa taaaacttcc aattatgatg caaaaaaaaaa aaaaaa 1379
130 <210> SEQ ID NO: 2
131 <211> LENGTH: 313
132 <212> TYPE: PRT

```

RAW SEQUENCE LISTING

DATE: 03/25/2005

PATENT APPLICATION: US/09/721,047A

TIME: 07:49:59

Input Set : A:\Ocirs396.app

Output Set: N:\CRF4\03252005\I721047A.raw

```

133 <213> ORGANISM: Phytolacca americana
135 <400> SEQUENCE: 2
136 Met Lys Ser Met Leu Val Val Thr Ile Ser Ile Trp Leu Ile Leu Ala
137   1           5           10           15
139 Pro Thr Ser Thr Trp Ala Val Asn Thr Ile Ile Tyr Asn Val Gly Ser
140           20           25           30
142 Thr Thr Ile Ser Lys Tyr Ala Thr Phe Leu Asn Asp Leu Arg Asn Glu
143           35           40           45
145 Ala Lys Asp Pro Ser Leu Lys Cys Tyr Gly Ile Pro Met Leu Pro Asn
146           50           55           60
148 Thr Asn Thr Asn Pro Lys Tyr Val Leu Val Glu Leu Gln Gly Ser Asn
149   65           70           75           80
151 Lys Lys Thr Ile Thr Leu Met Leu Arg Arg Asn Asn Leu Tyr Val Met
152           85           90           95
154 Gly Tyr Ser Asp Pro Phe Glu Thr Asn Lys Cys Arg Tyr His Ile Phe
155           100          105          110
157 Asn Asp Ile Ser Gly Thr Glu Arg Gln Asp Val Glu Thr Thr Leu Cys
158           115          120          125
160 Pro Asn Ala Asn Ser Arg Val Ser Lys Asn Ile Asn Phe Asp Ser Arg
161   130          135          140
163 Tyr Pro Thr Leu Glu Ser Lys Ala Gly Val Lys Ser Arg Ser Gln Val
164 145          150          155          160
166 Gln Leu Gly Ile Gln Ile Leu Asp Ser Asn Ile Gly Lys Ile Ser Gly
167           165          170          175
169 Val Met Ser Phe Thr Glu Lys Thr Glu Ala Glu Phe Leu Leu Val Ala
170           180          185          190
172 Ile Gln Met Val Ser Glu Ala Ala Arg Phe Lys Tyr Ile Glu Asn Gln
173           195          200          205
175 Val Lys Thr Asn Phe Asn Arg Ala Phe Asn Pro Asn Pro Lys Val Leu
176   210          215          220
178 Asn Leu Gln Glu Thr Trp Gly Lys Ile Ser Thr Ala Ile His Asp Ala
179 225          230          235          240
181 Lys Asn Gly Val Leu Pro Lys Pro Leu Glu Leu Val Asp Ala Ser Gly
182           245          250          255
184 Ala Lys Trp Ile Val Leu Arg Val Asp Glu Ile Lys Pro Asp Val Ala
185           260          265          270
187 Leu Leu Asn Tyr Val Gly Gly Ser Cys Gln Thr Thr Tyr Asn Gln Asn
188           275          280          285
190 Ala Met Phe Pro Gln Leu Ile Met Ser Thr Tyr Tyr Asn Tyr Met Val
191   290          295          300
193 Asn Leu Gly Asp Leu Phe Glu Gly Phe
194 305          310
197 <210> SEQ ID NO: 3
198 <211> LENGTH: 934
199 <212> TYPE: DNA
200 <213> ORGANISM: Phytolacca americana
202 <220> FEATURE:
203 <221> NAME/KEY: sig_peptide
204 <222> LOCATION: (1)..(75)

```

RAW SEQUENCE LISTING

DATE: 03/25/2005

PATENT APPLICATION: US/09/721,047A

TIME: 07:49:59

Input Set : A:\Ocirs396.app

Output Set: N:\CRF4\03252005\I721047A.raw

```

206 <220> FEATURE:
207 <221> NAME/KEY: mat_peptide
208 <222> LOCATION: (76)..(930)
210 <220> FEATURE:
211 <221> NAME/KEY: CDS
212 <222> LOCATION: (1)..(930)
214 <400> SEQUENCE: 3
215 atg aag atg aag gtg tta gaa gta gtt ggg ttg gca ata tcg ata tgg 48
216 Met Lys Met Lys Val Leu Glu Val Val Gly Leu Ala Ile Ser Ile Trp
217 -25 -20 -15 -10
219 ctg atg ctt aca cca cca gct tct tca aac ata gtg ttt gac gtt gag 96
220 Leu Met Leu Thr Pro Pro Ala Ser Ser Asn Ile Val Phe Asp Val Glu
221 -5 -1 1 5
223 aat gcc aca cca gaa acc tac tct aat ttt ctg act agt ttg cga gaa 144
224 Asn Ala Thr Pro Glu Thr Tyr Ser Asn Phe Leu Thr Ser Leu Arg Glu
225 10 15 20
227 gct gtg aaa gac aag aaa ttg aca tgc cat gga atg ata atg gcc aca 192
228 Ala Val Lys Asp Lys Lys Leu Thr Cys His Gly Met Ile Met Ala Thr
229 25 30 35
231 acc ctc act gaa caa ccc aag tat gtg ttg gtt gac ctc aaa ttc gga 240
232 Thr Leu Thr Glu Gln Pro Lys Tyr Val Leu Val Asp Leu Lys Phe Gly
233 40 45 50 55
235 tct gga aca ttc aca tta gca atc aga agg gga aac tta tat ttg gag 288
236 Ser Gly Thr Phe Thr Leu Ala Ile Arg Arg Gly Asn Leu Tyr Leu Glu
237 60 65 70
239 ggc tat tct gac att tac aat gga aaa tgt cgt tat cgg atc ttc aag 336
240 Gly Tyr Ser Asp Ile Tyr Asn Gly Lys Cys Arg Tyr Arg Ile Phe Lys
241 75 80 85
243 gat tca gaa tcc gat gcc caa gag acc gtt tgc ccc ggg gac aaa agc 384
244 Asp Ser Glu Ser Asp Ala Gln Glu Thr Val Cys Pro Gly Asp Lys Ser
245 90 95 100
247 aag cct ggc act cag aat aat atc ccc tat gaa aag agt tac aaa ggg 432
248 Lys Pro Gly Thr Gln Asn Asn Ile Pro Tyr Glu Lys Ser Tyr Lys Gly
249 105 110 115
251 atg gaa tca aag ggt ggg gct aga act aaa tta ggg tta gga aag ata 480
252 Met Glu Ser Lys Gly Gly Ala Arg Thr Lys Leu Gly Leu Gly Lys Ile
253 120 125 130 135
255 aca ctc aag agt cga atg ggt aaa atc tac ggc aag gat gca acg gat 528
256 Thr Leu Lys Ser Arg Met Gly Lys Ile Tyr Gly Lys Asp Ala Thr Asp
257 140 145 150
259 cag aag cag tat caa aaa aat gag gct gaa ttt ctt ctt ata gcc gtt 576
260 Gln Lys Gln Tyr Gln Lys Asn Glu Ala Glu Phe Leu Leu Ile Ala Val
261 155 160 165
263 caa atg gtt act gag gca tca agg ttc aaa tac att gag aac aaa gtg 624
264 Gln Met Val Thr Glu Ala Ser Arg Phe Lys Tyr Ile Glu Asn Lys Val
265 170 175 180
267 aag gct aaa ttt gat gat gcc aat ggg tat cag cca gat cct aaa gct 672
268 Lys Ala Lys Phe Asp Asp Ala Asn Gly Tyr Gln Pro Asp Pro Lys Ala
269 185 190 195

```

RAW SEQUENCE LISTING

DATE: 03/25/2005

PATENT APPLICATION: US/09/721,047A

TIME: 07:49:59

Input Set : A:\Ocirs396.app

Output Set: N:\CRF4\03252005\I721047A.raw

```

271 att tcc cta gag aaa aat tgg gac agt gtt tct aag gtc att gca aaa 720
272 Ile Ser Leu Glu Lys Asn Trp Asp Ser Val Ser Lys Val Ile Ala Lys
273 200 205 210 215
275 gtt ggc acc tcc ggt gat agt act gtt act tta cct gga gac cta aaa 768
276 Val Gly Thr Ser Gly Asp Ser Thr Val Thr Leu Pro Gly Asp Leu Lys
277 220 225 230
279 gat gag aat aat aaa cct tgg act acg gcc acc atg aac gac ctt aag 816
280 Asp Glu Asn Asn Lys Pro Trp Thr Thr Ala Thr Met Asn Asp Leu Lys
281 235 240 245
283 aac gac att atg gca ctc cta acc cac gtt act tgc aag gtt aaa agt 864
284 Asn Asp Ile Met Ala Leu Leu Thr His Val Thr Cys Lys Val Lys Ser
285 250 255 260
287 tcc atg ttc cct gaa att atg tcc tat tat tat agg act agt att agt 912
288 Ser Met Phe Pro Glu Ile Met Ser Tyr Tyr Tyr Arg Thr Ser Ile Ser
289 265 270 275
291 aac ctt ggt gaa ttc gag tgat 934
292 Asn Leu Gly Glu Phe Glu
293 280 285
297 <210> SEQ ID NO: 4
298 <211> LENGTH: 310
299 <212> TYPE: PRT
300 <213> ORGANISM: Phytolacca americana
302 <400> SEQUENCE: 4
303 Met Lys Met Lys Val Leu Glu Val Val Gly Leu Ala Ile Ser Ile Trp
304 -25 -20 -15 -10
306 Leu Met Leu Thr Pro Pro Ala Ser Ser Asn Ile Val Phe Asp Val Glu
307 -5 -1 1 5
309 Asn Ala Thr Pro Glu Thr Tyr Ser Asn Phe Leu Thr Ser Leu Arg Glu
310 10 15 20
312 Ala Val Lys Asp Lys Lys Leu Thr Cys His Gly Met Ile Met Ala Thr
313 25 30 35
315 Thr Leu Thr Glu Gln Pro Lys Tyr Val Leu Val Asp Leu Lys Phe Gly
316 40 45 50 55
318 Ser Gly Thr Phe Thr Leu Ala Ile Arg Arg Gly Asn Leu Tyr Leu Glu
319 60 65 70
321 Gly Tyr Ser Asp Ile Tyr Asn Gly Lys Cys Arg Tyr Arg Ile Phe Lys
322 75 80 85
324 Asp Ser Glu Ser Asp Ala Gln Glu Thr Val Cys Pro Gly Asp Lys Ser
325 90 95 100
327 Lys Pro Gly Thr Gln Asn Asn Ile Pro Tyr Glu Lys Ser Tyr Lys Gly
328 105 110 115
330 Met Glu Ser Lys Gly Gly Ala Arg Thr Lys Leu Gly Leu Gly Lys Ile
331 120 125 130 135
333 Thr Leu Lys Ser Arg Met Gly Lys Ile Tyr Gly Lys Asp Ala Thr Asp
334 140 145 150
336 Gln Lys Gln Tyr Gln Lys Asn Glu Ala Glu Phe Leu Leu Ile Ala Val
337 155 160 165
339 Gln Met Val Thr Glu Ala Ser Arg Phe Lys Tyr Ile Glu Asn Lys Val
340 170 175 180

```

VERIFICATION SUMMARY

DATE: 03/25/2005

PATENT APPLICATION: US/09/721,047A

TIME: 07:50:00

Input Set : A:\Ocirs396.app

Output Set: N:\CRF4\03252005\I721047A.raw